

EXPRESS MAIL LABEL NO.: EK873466237US DATE OF DEPOSIT: February 13, 2001
I hereby certify that this paper and fee are being deposited with the United States Postal Service Express Mail Post Office to Addressee service under 37 CFR \$1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents, Washington, D.C. 20231
Dianne Lane
NAME OF PERSON MAILING PAPER AND FEE SIGNATURE OF PERSON MAILING PAPER AND FEE

INVENTORS: Renee M. Kovales, Edith H. Stern, Barry E. Willner

Recording and Receiving Voice Mail with Freeform Bookmarks

BACKGROUND OF THE INVENTION

Related Inventions

The present invention is related to the following commonly-assigned U. S. Patents, both of which were filed concurrently herewith and are hereby incorporated herein by reference: U. S. _____ (serial number 09/_____), entitled "Selectable Audio and Mixed Background Sound for Voice Messaging System", and U. S. _____ (serial number 09/_____), entitled "Audio Renderings for Expressing Non-Audio Nuances".

Field of the Invention

The present invention relates to a computer system, and deals more particularly with methods, systems, computer program products, and methods of doing business by providing bookmarks for voice mail messages in order to facilitate improved navigation and processing thereof.

Description of the Related Art

Face-to-face communication between people involves many parallel communication paths. We derive information from body language, from words, from intonation, from facial expressions, from the distance between our bodies, and so forth. Distance communication, such as phone calls, e-mail exchange, and voice mail, on the other hand, involves only a few of these communication paths. Users may therefore have to take extra actions (which may or may not be successful) if they wish to try to overcome the limitations so imposed.

Distance communicating is becoming more prevalent in our society. Voice mail systems became widely used in years past, and in more recent years electronic mail systems have become common, with the popularity and pervasiveness of e-mail continuing to grow. When communicating by e-mail, message creators often try to overcome the limitations of distance communications by techniques such as using different font sizes, colors, emoticons (i.e. combinations of text symbols which bear a resemblance to facial expressions), and so forth to express non-text information. This non-text information includes emphasis, emotion, irony, etc. When communicating in person, the speaker can use changes in body language to indicate a

change in subject. In e-mail messages, the paragraph structure and use of bolding and italics gives clues as to the number and importance of topics. Thus, e-mail users try to overcome the limitations of distance communications by using visual clues for both semantic and contextual meaning. Further, visual clues such as paragraph distinctions are perceivable as soon as the page is displayed, providing a "broadside" perception of the message.

Voice mail has a different set of problems. While the recipient has the benefit of the nuances available through voice, the recipient does not have the advantages of the other parallel forms of communication which are available in person. Unlike e-mail, with voice mail the recipient does not have the advantage of broadside perception of the message. Thus, in a voice mail message, it may be difficult for the listener to appreciate when one topic has ended and another has begun. Additionally, voice mail users listening to their messages from a telephone do not have the ability to navigate within the stored voice mail in a controlled fashion. The voice mail recipient is hampered in the retrieval of, and subsequent actions on, the message due to the inability to act on parts of the voice mail message, as can be done with e-mail messages. (While some existing voice mail systems allow the listener to speed up the message, or skip the message, none are known to the present inventors which allow the originator of the message to specify where the topics lie, or which enable use of such identified topics for actions by the listener.)

Accordingly, what is needed is a technique that alleviates these problems in distance communications, providing a more flexible and more productive way for people to communicate using voice mail messages.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a technique that alleviates disadvantages in distance communications.

Another object of the present invention is to provide this technique by enabling a more
5 flexible and more productive way for people to communicate using voice mail messages.

A further object of the present invention is to provide these advantages by augmenting a voice mail message with bookmarks.

Still another object of the present invention is to provide these advantages by using the bookmarks to navigate stored voice mail messages more efficiently.

10 Yet another object of the present invention is to provide these advantages by using the bookmarks to enable acting upon partial voice mail messages.

Yet another object of the present invention is to provide new methods of doing business, whereby enhanced voice mail systems can be provided to end-users, and/or features of existing systems can be improved.

Other objects and advantages of the present invention will be set forth in part in the description and in the drawings which follow and, in part, will be obvious from the description or may be learned by practice of the invention.

To achieve the foregoing objects, and in accordance with the purpose of the invention as
5 broadly described herein, in a first aspect the present invention provides methods, systems, computer program products, and methods of doing business by providing bookmarks in voice mail messages.

This technique comprises providing one or more bookmarks for a voice mail message left
10 by a caller in order to create a bookmarked message. It may also comprise creating the resulting bookmarked message, and playing the bookmarked message to a listener.

The bookmarks may be used for a number of different purposes, including but not limited
15 to: segmenting the voice mail message according to one or more topics; marking one or more segments as having one or more different degrees of importance (in which case a default degree of importance may be associated with segments not otherwise marked); identifying one or more special types of information in the voice mail message (such as a callback telephone number for the caller; a callback time and/or date to use when responding to the voice mail message); identifying the caller's name; providing a Uniform Resource Locator or an e-mail address; and providing a protection or confidentiality indicator for the message or some portion thereof.

5

The bookmarks may be used to navigate from one topic or segment of the voice mail message to another, and/or to perform actions selectively on one or more segments, where the actions include, but are not limited to: saving; deleting; forwarding; skipping; listening; or repeating playback of segments; and/or marking a selected segment as protected or confidential (for example, prior to forwarding the segment to another party).

The bookmarks may also be used to navigate among segments of the voice mail message having one or more of the different degrees of importance, and/or to perform actions (of the type just described) selectively on one or more of these segments.

10

Similarly, the bookmarks may be used to navigate to one or more of the special types of information in the voice mail message, and/or to selectively perform actions (of the type described above) on one or more of these special types of information. When one of the special types of information is a Uniform Resource Locator, then one of the actions may comprise establishing a connection to that URL.

15

In some embodiments, the bookmarks and the voice mail message may be stored separately; in other embodiments, they may be stored in an intermingled manner. (Or, selected types of the bookmarks may be separately stored, while others are intermingled.)

The number of bookmarked topics in the voice mail message may be announced to a listener. Similarly, the number of different degrees of importance (and/or the number of segments

in these different degrees of importance) may be announced, as may the number of special types of information.

Optionally, particular ones of the special types of information may be treated as required, in which case the caller may be prompted to provide input for each required one for which no bookmark is otherwise provided.

As another option, one or more audio cues may be associated with one or more of the bookmarks. In this case, the technique may further comprise playing the bookmarked message to a listener while also playing the associated audio cues along with respective portions (e.g. segments and/or special indicators) of the voice mail message which correspond to the bookmarks. The audio cues may be incorporated in with the respective portions of the voice mail message, or may be played in-line (i.e. between the portions).

The present invention will now be described with reference to the following drawings, in which like reference numbers denote the same element throughout.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a flow diagram illustrating an example of how a caller may enhance his or her voice mail message by providing bookmarks, according to preferred embodiments of the present invention;

Figs. 2A through 2D provide a flowchart illustrating logic that may be used to implement preferred embodiments of the voice mail playback features of the present invention;

Figures 3A through 3D are tables showing examples of command choices that may be provided to a listener for use in navigating the enhanced voice mail messages of the present invention;

Figure 4 provides a histogram illustrating message segmenting, as defined by the present invention; and

Figures 5A through 5G depict examples of data structures that may be used to facilitate implementation of preferred embodiments of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention improves distance communications which use voice mail messages, providing a more flexible and more efficient voice mail system. Use of freeform bookmarks (i.e. bookmarks that may be provided anywhere within a voice mail message) in voice mail messages as disclosed herein enables disambiguating topics during the voice messages. Further, it allows “skimming” of voice mail analogous to a reader skimming the contents of an e-mail message. The disclosed techniques enable rapid navigation of voice mail messages from an ordinary, non-screen telephone, and rapid retrieval of information such as phone numbers or special text from the midst of a voice mail message.

elect to forward only the previous segment or the upcoming segment to a third party; or to delete, save, or replay the segment; and so forth. Rather than selecting such actions while hearing a voice mail message segment, the listener may alternatively choose such actions selectively, using (e.g.) a numeric identifier of the applicable message segment(s).

5 A number of different embodiments of the present invention may be implemented using the teachings disclosed herein. Preferred ones of these embodiments, illustrating the advantageous features of the present invention, will now be described.

10 The present invention may be used to enhance voice mail messages and voice mail systems by explicitly identifying various types of information in a voice mail message. Speakers tend to be less precise in oral speech than they are in written communications. This can be inefficient when communicating using voice mail. When leaving a message, the present invention enables the speaker to disambiguate topics of his voice message. A “keyed alert sequence” (“KAS”), which preferably comprises one or more signals entered by the user from his telephone equipment, is used to signal to the VMS that a bookmark of the voice mail message should be created. For
15 example, the KAS signal may be generated by the caller pressing keys or buttons on his telephone to transmit dual-tone multi-frequency (DTMF) tones. Or, selections may be made in any other available manner. For example, the user may speak a command or command sequence that will be interpreted by a voice recognition process (which may be part of the caller’s equipment, or part of the receiving voice mail system) and used as a KAS to generate a bookmark. Alternatively, the
20 caller may have a telephone with a display screen (such as a screen phone or a Web-enabled

importance indicator with a segment of his message (e.g. by keying a sequence such as “#3” or “***” to mark the segment as being of importance level 3). Or, a single level of importance indicator may be supported alternatively, where it is assumed that all of the spoken message is of the same (default) importance until the speaker provides a KAS to distinguish a particular passage. As an example of using this approach, a voice command may be given such as “important passage begin” as the caller begins speaking an important passage, and then “important passage end” may be spoken after the important passage has been communicated to the VMS.

Another example of advantageous use of KAS indicators as voice mail bookmarks is in highlighting special information (such as a phone number or an identifier the listener may need to record, or an address such as an e-mail address or a Uniform Resource Locator or “URL”) in a voice mail message. By bookmarking the information, the listener can navigate directly to the information he needs or can have it played back to him, without having to hear the entire message again. The types of special information supported may be predefined in a particular implementation, or a number of different indicators may alternatively be provided without having a fixed interpretation thereof. In the former case, the implementation may be adapted to storing a callback phone number after the caller provides a KAS such as “*1” and a callback time and/or date after the caller presses “*2”, or a URL when the caller presses “#URL”, for example. (A “callback” phone number, as the term is used herein, is a phone number which this caller wishes the listener to use when responding to this voice mail message, and which may differ from the number of the telephone from which the caller is leaving her message.) This approach enables the

10970827P.02404
10/01/2020 22:28:50

VMS to explicitly inform the listener of what types of special information have been stored for a particular voice mail message, as well as enabling separate rendering of that information to the listener. Furthermore, the VMS may optionally be adapted to know which of the special topics (if any) is considered required or preferred information, and may prompt the caller to provide this information. When special indicators are provided without a fixed interpretation, the semantic interpretation of the special indicators may perhaps be known only to the caller and the listener, with the VMS enabling announcement, separate playback, and/or navigation to the special information requested by the user. If URLs are supported as a type of special indicator, embodiments of the present invention may optionally provide a feature for automatically establishing a connection to that URL for the listener.

An implementation of the present invention may use KAS indicators for creating voice mail bookmarks for any of the above-described functions (as well as functions which will be obvious once the teachings disclosed herein are known) separately or in combination(s).

In one embodiment, the KAS indicator(s) may be recorded in-line as part of the audio stream (i.e. intermingled within the audio stream). In an alternative embodiment, KAS detection causes the development of metadata (such as an array of codes or pointers) which is associated with the audio stream. These techniques enable the VMS and listener to efficiently navigate the recorded message, and to perform actions selectively on portions of the message. Or, a combination of these techniques may be used. The storage of KAS indicators is described in more detail below, with reference to the sample data structures in Figs. 5A through 5G.

Reference is now made to Fig. 1, which provides an example that illustrates how a caller may enhance her voice mail message by providing one or more bookmarks to (1) indicate when the topic of her spoken message changes and (2) indicate when she is providing special information such as her callback phone number. As shown in Fig. 1, the caller 100 dials a number (which may be the called party's number, or perhaps a voice mail service to which the called party subscribes), as shown at 102. The call is answered 103 by the VMS 101, and the message recording starts 104. Preferably, the caller begins speaking 105 and provides 106 a KAS at some (arbitrary) call-specific point within the call. (Alternatively, the caller may choose to provide a KAS at the beginning of the call to mark the beginning of her first topic before speaking.) The character "*" is shown in Fig. 1 as the "change topic" KAS, for purposes of illustration and not of limitation. After receiving the KAS which the caller provided at 106, the VMS internally identifies 107 that a new topic has started for this voice mail message. This internal identification may comprise segmenting the voice mail message such that a change-topic bookmark can be associated with the new segment in metadata. Or, the KAS or an interpretation thereof may be inserted in-line within the stored message. (For example, a phrase "new topic" may be programmatically spoken into the recorded message. This programmatic insertion may be performed by the VMS, or by the caller's phone equipment or even by the telephone company as the message is being transmitted to the VMS.) In the preferred embodiment, an entry is created in metadata associated with the message. A data structure of the type depicted in Figs. 5A - 5G may be used for these approaches. (Note that if DTMF tones are used for providing the KAS, the receiving VMS preferably does not store the sound of the DTMF tones for the called party to hear, but rather interprets them and may store them for further programmatic processing.)

100
150
The caller then continues speaking the new topic of her voice message (108), which the VMS continues to record. For purposes of illustration, it is assumed that the VMS in Fig. 1 supports both change-topic bookmarks and special indicators, and that this user wishes to explicitly indicate information for special treatment, such as a phone number. At some point during her spoken message, the caller thus provides a KAS such as “#”, as shown at 109, to indicate that she is now providing a particular type of special information (indicated in the example as a phone number). In some embodiments, a special or first character such as “#” may indicate that special information is being entered, while an additional character or characters is then provided (e.g. comprising a sequence such as “#11”) to identify which special indicator this is. The VMS receives the KAS indicator, and internally identifies 110 that a phone number is about to be entered. The caller then keys in or speaks the phone number 111, after which she preferably provides another KAS – which may be a special KAS defined to signal the VMS of the end of the special information, or which may be a new topic indicator (or another special indicator), as shown in the example at 112. (Alternatively, in some cases the VMS may be adapted to know when the special information is complete. For example, provision of a fixed-length special message may be supported. In such cases, it is not necessary to provide an explicit KAS to signify the end of the special information.)

Upon receiving the change-topic KAS provided at 112, the VMS internally identifies 113 the start of the new topic, and the caller begins speaking 114 the information for that new topic.

This process of changing topics and/or providing special information may be repeated a number of

times, if desired. When the caller's message is complete, she hangs up 115, and the VMS stops recording this voice mail message, as shown at 116.

Note that while the example scenario provided in Fig. 1 refers to use of special symbols and numbers from the caller's telephone, this is for purposes of illustration and not of limitation.

5 Alternatively, any suitable technique for supplying input from the caller, including use of voice recognition by the receiving VMS to interpret the caller's spoken selection(s) or command(s), may be used instead. Furthermore, a particular implementation of the present invention may choose to support a combination of techniques for input.

10 While not illustrated in Fig. 1, a particular receiving VMS may choose to prompt a caller to provide special information such as the callback number or the caller's name. In this case, the prompting may occur by a spoken request from the VMS, or by transmitting indicators of another form. Optionally, at 107, 110, and 113, the VMS may provide confirmation to the caller that the KAS tones have been received. This confirmation may be a tone, spoken message, visual display, or other indication. It will be obvious to one of skill in the art how Fig. 1 may be adapted to
15 support either or both of these alternative approaches.

Figs. 2A - 2D provide a flowchart illustrating logic that may be used to implement preferred embodiments of the voice mail playback features of the present invention. This logic is invoked after the listener has already initiated interactions with the enhanced VMS and acts to receive a particular stored voice mail message for this listener, which the VMS has located. The

VMS may either proceed through each voice mail message for this particular listener sequentially, or may provide a means for the listener to request processing of individual messages, using prior art techniques. In either case, the features of the present invention allow the listener to process his voice mail messages more quickly and more efficiently. (Note that in the latter case, it is assumed that the means for selecting a message provides only an announcement of message sequence numbers, or perhaps of the phone number from which the message was delivered and its registered owner. The ability to skim the actual content of the messages by subject or to navigate messages based on topics or other special indicators, as disclosed herein, is beyond the scope of the prior art techniques.)

The logic in Fig. 2 assumes that the enhanced VMS supports bookmarks that provide for message segmenting (such as the change-topic bookmarks described above) as well as special tagged fields (i.e. the special indicator types which have been described, such as callback numbers). This logic therefore begins the processing of a voice mail message at Block 200 with the VMS announcing the number of segments and the number of tagged fields which have been stored for this particular message. The announcement may be spoken to the listener, or conveyed in another manner (such as transmitting a numerical indicator or other value for display on the listener's cell phone or other equipment). This announcement gives the listener a sense of the topics to be covered, in a similar manner to how a person viewing her e-mail gets a sense of the message by seeing the paragraph layout. In some preferred embodiments, an array or similar data structure is used for storing bookmarks. In these embodiments, the number of segments and tagged fields may be determined, for example, using a software function that obtains the number

of elements in the array. In other embodiments, the segment markers and tags may appear in-line within the message, in which case the VMS preferably processes the message to count the number of segments and tags. (This counting may be done as the message is being spoken and recorded, or in a batch mode after the message is recorded, or as the message is accessed for playback.)

5 At Block 202, the VMS prompts the listener to see what action he would like to take for this stored message. Block 204 then checks to see which option was selected. A representative set of options is shown in the table in Fig. 3A. As shown therein, the listener may choose to listen to the entire message (as in a prior art VMS); listen to a particular segment; access the tagged fields for the message; or exit. Representative command sequences are also shown in Fig. 3A, by way of example, that may be used in response to the prompting of the VMS. (As will be obvious, the selections shown in Figs. 3A through 3D, as well as the commands that may be used to request each of these selections, are intended merely as examples.)

10 If the listener chooses the exit option (using the response "99" in the example), then the processing in Fig. 2A for this voice mail message ends. If the listener chooses to hear a particular
15 segment of the message (using the response "#n", where the value of "n" is a number corresponding to the requested segment), control transfers to Block 206 where this "n-th" segment is played. The listener is then prompted (Block 208) for his next action on this segment. Control then transfers to Block 220 of Fig. 2B, where a test is made to see which action the listener requested for this segment. Representative choices are shown in Fig. 3B. If the listener
20 chooses to delete the segment (option 1), then at Block 224 that segment is deleted from the

stored message. If the listener chooses to save the segment (option 2), then at Block 226 that segment is saved (for example, for further reference or further processing). If the listener chooses to forward the segment (option 4), then at Block 222 that segment is forwarded to a third party. (The destination phone number to use for this purpose is preferably obtained from the listener or from other means, using techniques which do not form part of the present invention and which are not shown in Fig. 2B.) After operation of Blocks 222, 224, or 226, or when the listener selects to continue (option 3), control returns to Block 202 of Fig. 2A to enable the listener to continue processing this voice mail message. (Alternatively, an implementation of the present invention may enable control to return to Block 208 of Fig. 2A after Blocks 222, 224, or 226, in order to accept additional segment-related requests from the listener; in this case, selecting "continue" enables the listener to return to the mainline processing at Block 202.)

Returning now to Block 204 of Fig. 2A, if the listener chooses to access a particular tagged field of the message (using the response "*m", where the value of "m" is a number corresponding to the requested tagged field), then at Block 210 he is prompted for the particular type of action he would like to perform on this "m-th" tagged field. The table in Fig. 3D provides representative examples of selections that may be provided. Control then transfers to Block 230 of Fig. 2C, which checks to see which option the listener chose. If he chooses to listen to the tagged field (option 1), then its contents are played (Block 234). If he chooses to forward the contents of this tagged field (option 4), then the contents are forwarded (Block 232) to a third party. (See the discussion of a third party above, with reference to Block 222 of Fig. 2B.) Or, other appropriate actions may be provided, as shown at Block 238 (in response to option 2). The

type of other actions, and the manner of carrying out such other actions, may depend on the type of tagged fields supported in a particular implementation.

One use of tagged fields may be to mark certain message content as being confidential or otherwise protected. When this feature is used, the caller preferably identifies the content to be protected at the time of leaving (i.e. speaking) that part of the voice mail message. In this case, it may be desirable to inhibit the listener's use of forwarding for this segment (not illustrated in the segment processing logic of Fig. 2C). Optionally, an implementation of the present invention may provide a similar protection function to be used by the listener, where (for example) the listener may save or mark message segments and/or tagged fields with "do not forward" attributes. This type of special features may be supported for use by callers or listeners, or both if desired. Furthermore, when supported, the feature may be implemented as a (standalone) special tagged field, or it may be supported as an additional indicator that may be associated with an entire message, with individual message segments, and/or with other tagged fields. It will be obvious to one of skill in the art how these optional features may be added to the logic of the Fig. 2. (Note that providing functionality to enforce such functions as "do not forward" is optional. Such functionality may be added to the present invention using techniques which are known in the art.)

After operation of Blocks 232, 234, or 236, or when the listener selects to continue (option 3), control returns to Block 202 of Fig. 2A to enable the listener to continue processing this voice mail message. (Alternatively, an implementation of the present invention may enable control to return to Block 210 of Fig. 2A after Blocks 232, 234, or 236, in order to accept

additional tagged field requests from the listener; in this case, selecting “continue” enables the listener to return to the mainline processing at Block 202.)

Referring again to Block 204 of Fig. 2A, if the listener chooses to play the entire message (using response “#0”), then control transfers to Block 212 where the message is played in its entirety. When tagged field bookmarks and/or topic change bookmarks are embedded in-line in a stored message, then this playback preferably comprises providing the listener with an audible indication of each bookmark. For example, a spoken phrase “topic is changing” may be inserted into the caller’s voice message, or a phrase such as “callback telephone number follows” may be inserted. Similarly, when bookmarks are stored as metadata, the metadata processing may include providing an audible indication, announcement, or visual indication during playback of the message. As an alternative to (or in addition to) audibly announcing the bookmark, audio cues may be incorporated in with the message playback (e.g. as background sound), where a change in the audio cue signals to the listener that a bookmarked field is being played. Use of audio cues with the present invention is discussed in more detail below.

Upon completing the playback, the listener is prompted for his next action (Block 214), after which control reaches Block 240 of Fig. 2D. The table in Fig. 3C provides representative examples of the actions that may be performed on complete messages. Block 240 checks to see what next action was selected. If the listener chooses to delete the entire message (option 1), then it is deleted (Block 244). If he chooses to forward the message (option 4), then the entire message is forwarded (Block 242) to a third party. (See the discussion of a third party above,

with reference to Block 222 of Fig. 2B.) If the listener chooses to save the entire message (option 2), then it is saved (Block 246) for later processing or other purposes. After operation of Blocks 242, 244, or 246, or when the listener selects to continue (option 3), control returns to Block 202 of Fig. 2A to enable the listener to continue processing this voice mail message. Alternatively, an implementation of the present invention may enable control to return to Block 214 of Fig. 2A after Blocks 242, 244, or 246, in order to accept additional requests from the listener which apply to the entire message; in this case, selecting "continue" enables the listener to return to the mainline processing at Block 202.)

The listener may iterate through the logic of Figs. 2A - 2D a number of times, if desired. Upon finishing with this message, the listener preferably chooses a selection such as "99", as stated above. This selection may signal the VMS to proceed to the listener's next stored message, or (if this is the last message) to stop processing messages for this listener. A particular implementation of the present invention may also provide a choice that enables the listener to directly exit the voice mail system at this point; or, a higher-level menu may be used to provide that option.

In another preferred embodiment not illustrated by the figures, the listener may be permitted to play back the message, and barge in with control requests. That is, even though the VMS may not prompt the listener for command responses, the listener may interrupt the message at any point by keying a code (e.g. *) and either in response to a prompt, or without further prompting, select an action to be taken. Such actions may include skipping the current segment,

deleting the current segment, proceeding to the next confidential segment, or any other command provided by the VMS. A barge-in feature is well known in the art of VMS, and is used in existing systems to perform such actions as skipping to the end of a message.

If KAS indicators for degrees of importance are supported for message segments, then the VMS may prompt the listener as to whether he would like to listen to (or act upon) segments of one particular importance level. Furthermore, the VMS may announce the degrees which have been used for the segments of a particular message. The listener may choose to listen to only the most important segments of each message or to segments of particular importance levels and so forth. While these features have not been illustrated in Fig. 2, it will be obvious to one of skill in the art how Fig. 2 may be adapted for this purpose.

Turning now to Fig. 4, a histogram is depicted which indicates graphically the effect of segmenting a voice mail message. In this example, a stored voice mail message has been divided into 5 segments, and each segment has some unique duration of message playback. Rather than listening to the entire message in sequence, as in the prior art, individual ones of the segments may be played back (or otherwise accessed or acted upon), providing a more flexible and more efficient technique for working with voice mail messages than is available in the prior art.

Figs. 5A through 5G depict examples of data structures that may be used by an implementation of the present invention. As will be obvious, these are merely one format in which the information may be stored. (In particular, while the data structures are shown as using arrays,

this is for purposes of illustration only; alternatively, linked lists or other similar structures may be used.)

Fig. 5A illustrates a data structure 500 that may be used to process and/or store voice mail segments which have tagged fields, where information about the tagged fields is stored separately from the message contents as metadata (while the contents of the tagged fields are preferably stored in-line within the message). In preferred embodiments, the tagged fields are stored in-line within the message as well. (In this manner, the tagged fields may be efficiently accessed during an in-order playback of the message, and also may be very quickly retrieved by indexing the separate array if necessary – e.g. if the listener requests to perform actions selectively using the tagged fields.) Each message may comprise some number of message segments, where the message is stored in “vmail_array” and the segments are stored as elements of the array (see 504). The contents of “vmail_array” (as well as the information in the structures of Figs. 5B through 5G) will preferably be stored in temporary storage while processing a particular message and accessing its stored elements using constructs such as the pointers shown in the figures. In addition, information for each message is preferably stored in some persistent data store such as a database, file structure on a disk, flash, etc. (When storing this information in persistent storage, the pointers to files are preferably replaced by file names or other indicators of the location of the referenced information.)

As shown at 503, each “vmail_segment” preferably comprises (1) a pointer to a location (such as a file) in which the data for this segment of the voice mail message is stored

10 (alternatively, the content could be stored in-line, rather than in a separate location), and (2) an
array of tagged field information. The pointer may point to a file stored in “.wav” format, or any
other appropriate format. The tagged field information, as shown at 502, preferably comprises
(1) a type indicator, which may be numeric, to indicate whether this is a phone number, an e-mail
5 address which the caller has provided for responding to the message or which pertains to the
message content, a URL, etc., and (2) a pointer to a location where the contents of the tagged
field are stored (alternatively, the content could be stored in-line, rather than in a separate
location). An upper limit on the size of the tagged field array and the segment array is shown at
501 for purposes of illustration.

10 The segments of a message may be processed (for example, during playback or when
preprocessing for later playback) using an in-order traversal of the arrays shown in the data
structures of Figs. 5A through 5G. Or, the arrays used to store information for each segment may
be indexed to selectively locate a particular segment.

15 Fig. 5B illustrates a data structure 510 that may be used to process and/or store voice mail
segments which have tagged fields, where information about the tagged fields is stored separately
from the message as metadata but the tagged field contents are stored in the message. In this
example, the tagged field contents are accessed using an offset calculation approach. This
structure 510 is identical to the structure 500 of Fig. 5A, except in the definition of “tagged_field”
512. In the approach of Fig. 5B, each tagged field is stored with (1) a type indicator, as described
20 for Fig. 5A, (2) an offset specifying the location within the “data_loc” of “vmail_segment” 513

where the data for this tagged field begins, and (3) a length that is used along with the starting offset to determine where the data for the tagged field ends.

The data structure 520 illustrated in Fig. 5C may be used when bookmarks are stored in-line within the voice mail message (perhaps as a spoken interpretation of the bookmark, for example, as discussed above). Preferably, this structure 520 comprises a pointer to a location (see 521), such as a file on disk, where all of the content for a particular message is stored. This file would then be searched to determine the segments, tagged fields, and so forth.

The data structure 530 illustrated in Fig. 5D may be used when bookmarks are used to segment or tag voice mail messages, and pointers to those segments or tags are stored in an array. If an indicator of the bookmark type (which is preferably stored in-line) is also stored in the message, then an interpretation of the bookmark may be provided (such as an announcement to the listener); otherwise, when the bookmark type is not stored in the message with each segment then semantic interpretation of the indicator type is not known (as discussed above). Preferably, this structure 530 comprises an array of a fixed maximum length (see 531 and 532), which is stored in temporary or persistent storage of some type, where the array elements comprise pointers to files which contain the data for each identified part of the voice mail message (or pointers to locations within a single file, which has not been illustrated in Fig. 5D).

Fig. 5E illustrates a data structure 540 which may be used for voice mail messages which have both tagged fields and separate indicators of special types of tagged fields. Each voice mail

message comprises an instance of "vmail_msg_struct" 544, which in this example contains separate pointers to locations where the special tagged fields (i.e. the caller's name and callback phone number) are stored, and an embedded array structure used for processing the message contents. This approach assumes that the contents of the special tagged fields have been explicitly
5 obtained or identified (or, alternatively, the pointers may be set to null to indicate an absence of such information); this information may also appear in-line within the message content.

Additional special tagged fields may be similarly stored in an analogous manner (or special tagged fields may alternatively be stored in a similar manner to the other tagged fields). Each segment of the message preferably uses the structure "vmail_segment" 543, which comprises a pointer to a
10 location where the structure's content is stored and an array of tagged field information (where the tagged field information is defined at 542). An upper bound on the array sizes is preferably fixed as shown at 541. Refer to Fig. 5A above for a discussion of the tagged field information and the "vmail_segment" information, which is used in an analogous manner for Fig. 5E.

Fig. 5F illustrates a data structure 550 that may be used for voice mail messages that have
15 special types of tagged fields identified, but do not have other tagged fields. As shown at 552, the data structure for such messages preferably comprises explicit pointers to the contents of each of the special tagged fields (in the example, the caller's name and callback phone number, as well as the desired callback time), along with a pointer to a file which stores an array of segments that
20 comprise the in-line content of the voice message. (This special tagged field information may also be stored in-line within the segments of the message.) An upper limit is preferably placed on the size of this array, as shown at 551.

Fig. 5G illustrates a data structure 560 that may be used for voice mail messages that have importance levels associated with segments as a special type of tagged field. As shown at 563, the data structure for such messages preferably comprises an array containing entries for each segment, where an upper limit shown at 561 is placed on the number of such entries. Each entry, shown at 562 using the structure "vmail_segment", preferably comprises an integer value for the importance level of this segment and a pointer to a location where the segment content is stored (although alternatively, the content could be stored in-line within the array element).

Note that the voice mail system may employ a database to contain the voice messages and/or their pertinent information, and the structure of each message and message segment as well as the element types used therein may therefore vary from what is shown in Figs. 5A through 5G – although the basic stored information is preferably the same.

According to an optional aspect of the present invention, audio cues can be used to provide additional contextual information while a segmented message or a message with tagged fields is being played to a listener. Such cues may be musical (tunes, for example). Or, they may be natural sounds, such as a birdsong, ocean waves, etc. As one example, audio cues can be used to indicate the degree of importance of the message segment. A background hum, incorporated in with the audio stream resulting from the translation, might indicate importance, with higher pitches indicating more important segments and lower pitches indicates less important. As another approach, the pitch or volume of the voice used for the audio rendering might change to indicate that the segment importance varies. Or, a background audio cue might change to a

completely different sound while message segments of different importance are being rendered.

Audio cues can also be provided for indicating information other than the relative importance of a message segment. As an example, the sound of a ringing phone might be used as an audio cue while a message segment tagged as providing a callback number is played to the listener. Or, the

5 VMS might associate different audio cues with each segment of a delineated message (which provides, for example, for playing a different audio cue per topic). As an alternative, the VMS may simply alternate between two audio cues, with adjacent message segments each having one of the cues. Upon playback of the message to the listener, the audio cues are preferably incorporated in with the spoken message. The association of the audio cues with sections of the message may be done as the message is retrieved for playback (in which case the VMS may perform the incorporating, or it may be done by the listener's telephone equipment, or perhaps by the telephone company); or, the association may be done prior to retrieval, for example by the VMS evaluating its stored messages to determine those which make use of audio cues and then processing those stored messages.

15 The related invention titled "Selectable Audio and Mixed Background Sound for Voice Messaging System" contains a thorough discussion of use of background sounds, and how such sounds may be selected, located, and processed. Refer to this related invention for more information about options and techniques that may be used for audio cues of the present invention. Note that this related invention is not directed toward inserting an audio cue or sound in-line as message content while a message is being rendered (e.g. a giggle sound in place of a smiley-face emoticon), although this in-line insertion approach for audio cues may be used with

the present invention to enhance voice mail messages. Or, the present invention may incorporate an audio cue as additional background sound for a voice mail message that is being rendered – or for some part of a message that is being rendered.

5 U. S. Patent 5,889,840, entitled “Voice Mail Replay Command System”, teaches a technique whereby a caller leaves a voice mail message, and the called party can embed answers within this message and return it to the caller. However, there is no teaching therein of the caller explicitly delineating message segments or of independently performing actions on partial voice mail messages.

10 U. S. Patent 5,983,187, entitled “Speech Data Storage Organizing System using Form Field Indicators” teaches a technique for capturing and storing speech data records according to “form field indicators”, which are elements provided by the speaker (such as a silence of a predetermined length, or pressing a button) to segment a voice data recording. Form field indicators correspond to information in the speech data (see column 4, lines 51 - 56), and serve as markers or anchors to be used when searching the speech data for a “form field”. A form field is a field corresponding to a form-oriented application. For example, a phone book application may comprise form fields of name, phone number, and comment. Using the disclosed technique, the user pauses or otherwise inserts a marker into his speech and speaks a keyword corresponding to one of a set of form fields. The system later processes the speech stream by locating the marker and then searching the spoken phonemes of the following speech data for any of a set of
15
20 predetermined keywords. Upon finding a phoneme that matches a keyword known to the system,

the speech information following that phoneme (up to the point where the next form field indicator is found) is stored into the corresponding form field. (See column 3, lines 13 - 25; lines 32 - 34; lines 43 - 45; and lines 53 - 64, as well as column 4, lines 12 - 20.) In this manner, if the user says "My phone number is 800-555-1212" and the phrase "phone number is" has been predefined as a keyword phrase, the system will detect the "phone number is" keyword phrase and store the spoken information "800-555-1212" which follows it into the phone book application's phone number field. Use of form field indicators is stated as being optional, provided that voice recognition is capable of detecting the phonemes of the spoken keyword phrase which introduces the form field content without use of an anchor. (See column 4, lines 44 - 46.) However, the disclosed techniques do not teach features of the present invention including performing actions on partial voice mail messages; navigation of voice mail messages by a listener wherein (for example) the listener skims from one bookmark to another; using bookmarks to highlight important passages within a voice mail message; announcing information about the stored content of a voice mail message; disambiguating topics using freeform bookmarks; or using audio cues as background sound to provide an audible representation of contextual information during playback to a listener.

U. S. Patent 5,742,736, entitled "Device for Managing Voice Data Automatically Linking Marked Message Segments to Corresponding Applications" teaches a technique for enabling a user to associate speech with a marker type, where the marker type is already associated with an application. A previously-stored voice message (or other type of voice data) is analyzed by a listener, using either a visual representation (e.g. a graphical display such as a line or bar) of a

voice message that is being spoken or an audio playback thereof. The user selects from among a set of predefined markers, and associates a marker with a portion of the speech (e.g. by dragging a marker icon in the visual display, or pressing keys while hearing a message played). An application corresponds to each marker, and the user's association automatically causes a linking between the marked speech and storage for the application. For example, if the user determines which part of a visual display corresponds to a recorded phone number, he may drag a phone icon to that segment, and the corresponding stored speech is then linked with the storage structure of a phone book application. In this manner, the stored speech is more easily integrated into existing applications. (See column 4, lines 38 - 54, as well as column 4, line 64 - column 5, line 3.) The present invention, on the other hand, is not directed toward marking a message while it is being played to the listener, but rather while it is being left by the speaker. In other words, the speaker creates the bookmarks of the present invention, while a listener marks or flags portions of messages in the prior art invention. Furthermore, the present invention does not require applications to be associated with bookmarks: the speaker will typically have no idea of the applications available to the listener, or of the applications with which the listener might like to associate the information from a voice mail message. Accordingly, the present invention is not directed at linking information from a voice mail message into the data structure used by an associated application.

U. S. Patent 5,526,407, entitled "Method and Apparatus for Managing Information", teaches a technique for recording, categorizing, organizing, managing, and retrieving written and speech information. This patent teaches an integrated system which provides a graphical user

interface to allow a personal computer or tablet user to manage audio data obtained from a meeting or other interaction. The disclosed technique is intended to allow a user (preferably, with a visual interface) to organize audio data; it does not teach bookmarking a voice mail message by the speaker.

5 As has been demonstrated, the present invention provides advantageous techniques to alleviate disadvantages of distance communication, for example by delineating topics or other significant information in voice mail messages, and for enabling listeners of voice mail messages to more efficiently navigate and process their messages. Once the teachings of the present invention are known, bookmarks in voice mail messages may be used advantageously in a myriad of ways to
10 enhance distance communications.

 As will be appreciated by one of skill in the art, embodiments of the present invention may be provided as methods, systems, or computer program products. Accordingly, the present invention may take the form of an entirely hardware embodiment, an entirely software
 embodiment or an embodiment combining software and hardware aspects. Furthermore, the
15 present invention may take the form of a computer program product which is embodied on one or more computer-usable storage media (including, but not limited to, disk storage, CD-ROM, optical storage, and so forth) having computer-usable program code embodied therein.

 The present invention has been described with reference to flowchart illustrations and/or flow diagrams of methods, apparatus (systems) and computer program products according to

embodiments of the invention. It will be understood that each block of the flowchart illustrations and/or flow diagrams, and combinations of blocks in the flowchart illustrations and/or flows in the flow diagrams, can be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general purpose computer, special purpose computer, embedded processor or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions specified in the flowchart and/or flow diagram block(s) or flow(s).

These computer program instructions may also be stored in a computer-readable memory that can direct a computer or other programmable data processing apparatus to function in a particular manner, such that the instructions stored in the computer-readable memory produce an article of manufacture including instruction means which implement the function specified in the flowchart and/or flow diagram block(s) or flow(s).

The computer program instructions may also be loaded onto a computer or other programmable data processing apparatus to cause a series of operational steps to be performed on the computer or other programmable apparatus to produce a computer implemented process such that the instructions which execute on the computer or other programmable apparatus provide steps for implementing the functions specified in the flowchart and/or flow diagram block(s) or flow(s). Furthermore, the instructions may be executed by more than one computer or data

processing apparatus, such that some aspects of the present invention operate (for example) at a VMS while other aspects operate the caller's or listener's telephone equipment, or elsewhere.

While preferred embodiments of the present invention have been described, additional variations and modifications in those embodiments may occur to those skilled in the art once they learn of the basic inventive concepts. Therefore, it is intended that the appended claims shall be construed to include both the preferred embodiments and all such variations and modifications as fall within the spirit and scope of the invention.

5

09/29/2014 10:39 AM
"09/29/2014 10:39 AM"